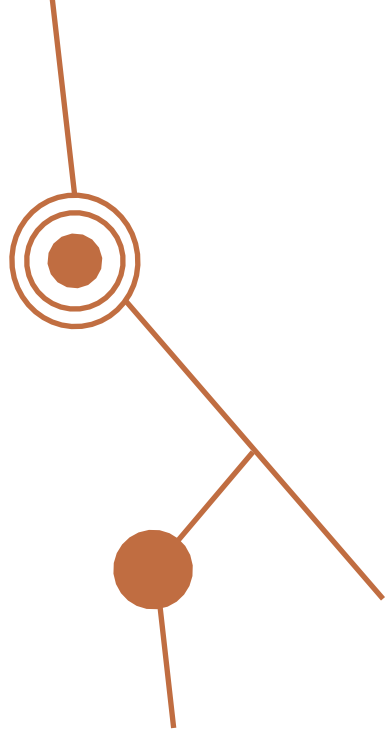
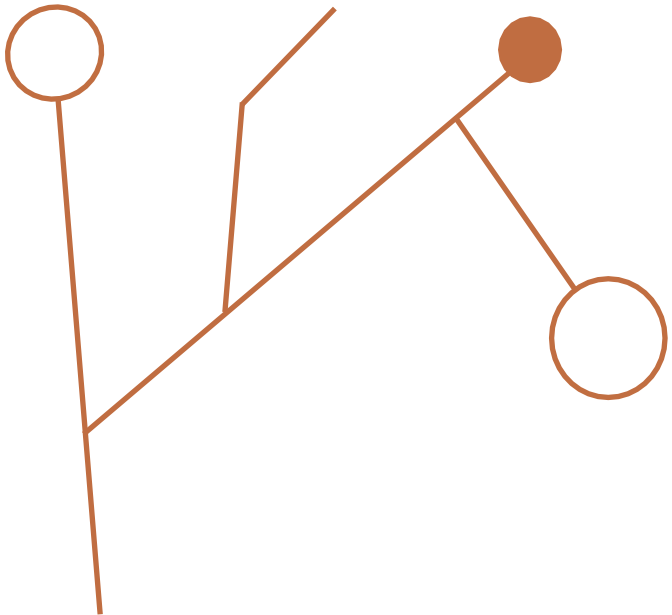
The future of care and algorithms

Van Voorst, R.S. 2023.

University of Amsterdam
European Research Council

Content





Introduction

This report presents the outcomes of an expert meeting on the future of healthcare and, more specifically, the increasing collaboration between medical experts and algorithms. Thirty participants took part in the roundtable, selected by the researchers for their extraordinary expertise and/or experience with algorithms in healthcare. The participants were aware of each other's diverse backgrounds, fostering an environment conducive to sharing knowledge and enriching one another's perspectives. Participants included professors of Artificial Intelligence, technology developers from relevant business sectors, futurists specialising in the healthcare field, an ethicist with expertise in healthcare and health insurance, code developers, social scientists and artists whose work focuses on the digitisation of healthcare. The expert meeting is part of a broader, multi-year study on human-non-human (algorithmic) collaboration in healthcare (see section 'about the study').

This report first describes the main themes and questions that emerged during the expert meeting. It then describes two future scenarios that participants indicated as realistic (but not necessarily desirable) for the near future.

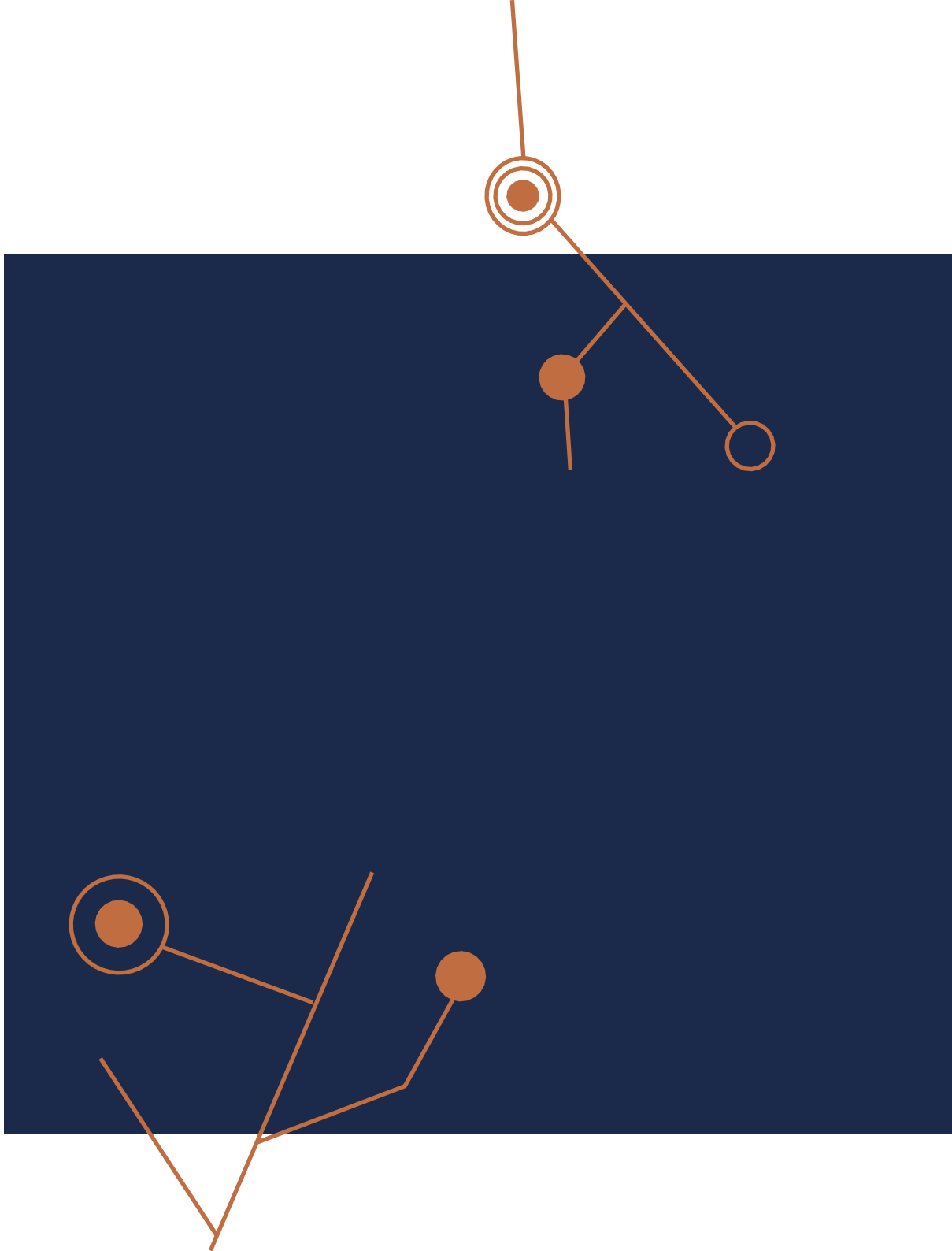
These scenarios are "ideal types" as the sociologist Max Weber described them:
constructed models used to approximate reality by selecting and highlighting certain elements.
Scenarios thus offer an outline of how the future of healthcare might unfold, based on developments that the participating experts see in their fields.

It is important to note that while these scenarios do not currently exist and it is not certain that they will materialise (as the future does not yet exist and will depend on the decisions and actions that policymakers and other stakeholders will take in the coming period), they are not entirely imaginary. Some aspects of them already occur in the present time.
A number of these are described in this report as example case studies.
To ensure the privacy of participants, these cases have been described without naming specific stakeholders or hospitals.

About the study

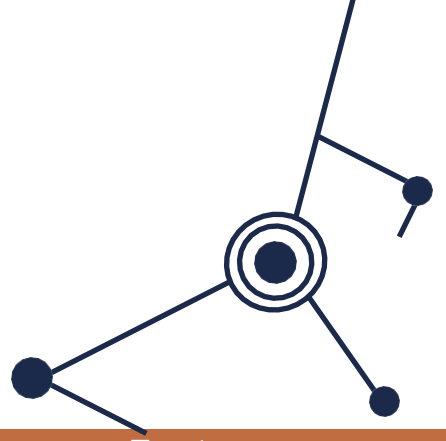
When people talk about 'algorithmic decision-making', they usually think of a computer system and the datasets it uses. In reality, all algorithmic decisions are made in collaboration with humans: we create them, evaluate them and apply or deviate from them. This research project is an anthropological study of the collaboration between humans and algorithmic systems in the field of global public health, a field where the growth in datafication and automation is unprecedented. In six countries, including the Netherlands, it examines how doctors, programmers and algorithms make decisions together, for example in DNA genetic testing or preventive healthcare. The research is led by future anthropologist Dr Roanne van Voorst and will be carried out by a team of social scientists between 2023-2028. The project is funded by the European Commission (ERC) and facilitated by the University of Amsterdam.

One of the aims of the study is to reflect together with experienced experts and other invested experts on what constitutes truly "fair" or "ethical" artificial intelligence, and how we can jointly work on this as stakeholders, as a society, doctors and developers. Empirical research in six case countries will show whether and how algorithms influence doctors' decisions, and what the effects (positive or negative) are on the daily work of doctors and public health.
In addition, surveys, roundtables, workshops and interviews with experts will lead to interpretations of future scenarios that will then be 'recreated' in a 'living' format in, for example, a theatre play with professional actors, a museum exhibition based on artist's creations or online visualised training for professionals. The idea behind this is that people can only understand a potential future scenario (and decide whether they find that future desirable or not) if they can experience it. This report is a first step in that direction.



Central themes and questions

Below is a brief overview of the central themes and questions that were discussed during the expert meeting. Some of these are then further developed into a scenario.



Responsibility

Is potentially effective healthcare technology adopted without clear accountability and without testing it in daily practice? Or is it strictly framed first, leading to delays? Do medics need to become "digitally literate" so that they know exactly how algorithms work, as is often claimed by policymakers? How do they integrate this new knowledge into their already heavily burdened tasks and who supports them in doing so, and is this requirement actually realistic or desirable? Who supervises when medical experts collaborate with an algorithm or its developers? Who is responsible if the technology makes mistakes in a medical environment?

Humanity

How can we ensure non-verbal communication when people increasingly interact with automated systems? How do doctors maintain their intuitive sense, which they develop throughout their careers? How do we teach computers ethics, or do we need to remind doctors that computers are essentially unethical? What happens to the development of human knowledge and skills if less attention is paid to this aspect in medical training (in exchange for a stronger technical focus)? Are human skills lost with the advent of A.I.? Can 'off-the-self ethics' contribute to an ethical relationship between doctor and patient? How exactly can Artificial Intelligence and other algorithmic systems be used to preserve the human aspect of medical work? Can we improve efficiency using A.I. as a tool? How do we prevent new technology from creating 'shadow work' for doctors, meaning extra work needed to make the technology work effectively?

Trust

Do doctors trust algorithms? Do patients trust algorithms? How can this trust be built? Is it desirable as currently stated in policy guidelines on the future of healthcare, or should doctors and patients instead be encouraged to remain highly sceptical of technology? Do we actually know enough about how people's decision-making is influenced by computer systems they interact with? (See also "human-non-human collaboration, who decides, what?") How will doctors be prepared for what is to come in the future in, for example, medical education? How will patients be prepared for these changes?

Central themes and questions

Education/work

What will medical education look like in the future? Is there already, and soon certainly, sufficient attention to preparing for human-technology interaction? Should training focus more on skills that cannot be taken over by computers? What competences do doctors, training and patients need in the process with A.I. as a key component? Should algorithms in healthcare mainly contribute to existing tasks/processes, or take over entire tasks/processes? Is enough attention being paid to "shadow work" or foreseeable "blind spots" of human-non-human interaction in training? Can artificial intelligence be used to improve doctors' work (e.g. more time for actual conversations with patients, as the algorithm takes over standard tasks, or some of the bureaucracy)? Can algorithms be used for simple, repetitive actions? Who will learn to control these actions? Can an automated system also diagnose or draw conclusions?

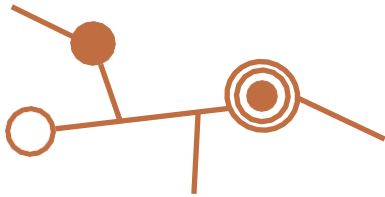
Ownership

Who will regulate A.I. and other forms of algorithms in healthcare, and who will be allowed to share or store data? How great is the threat of commercialisation to privacy and ownership? Who retains ownership of the patient database obtained or used? Who is allowed to earn from this data? Will patients be compensated for their data share?

Human-non-human cooperation: who decides, what?

Within the frameworks, guidelines and laws on artificial intelligence that are currently being developed, including for healthcare, it is invariably emphasised that algorithms always need a "human-in-the-loop" to work ethically: a human who can intervene if the algorithm gives remarkable, risky or wrong advice. While that sounds logical, it is not yet clear whether people are actually inclined to deal with their non-human, algorithmic colleagues in that assertive, controlling way. Research shows that people often do not stand up to the algorithm because, for example, they do not know exactly how it works or do not have time to check it. Sometimes the workload even increases after technology is introduced because more efficiency is expected from employees. Research on US judges, for example, showed that they often rely on an algorithm's advice because they do not have time to check it. Less experienced dermatologists were found to sometimes mistakenly rely on an algorithm's advice, while more experienced dermatologists more often reject the algorithm's advice and prefer to use their own insight and experience. The research project from which this report emerged aims to better understand when doctors do or do not rely on an algorithm and how this process can be designed more safely.

Scenario 1 Adoption of human tasks and processes



In 2033 A.I. is increasingly deployed with the aim of taking over tasks and processes in healthcare independently. The technology offers various opportunities in terms of expertise and efficiency, with the aim of helping patients faster, better and cheaper.

Where doctors are generalist experts, A.I. can learn new knowledge quickly and specifically. A.I. can quickly and accurately recognise patterns and then perform analysis. This offers interesting opportunities, for instance to efficiently analyse scans and skin blemishes. Increasingly, A.I. is therefore chosen when dealing with complex cases. However, this does carry an increased risk (not only because of potential 'black box' issues or other known problems surrounding A.I.; even a seemingly simple algorithm that simply follows a stylistic protocol can potentially influence a doctor's decisions as shown in: Human-non-human collaboration: who decides, what?).

This radically changes the role of medics. They not only need more screen time in their daily work, but also other tasks: they have to constantly verify and check the knowledge and data that A.I. provides. It is true that A.I. has a better prognostic value than a doctor, but the doctor’s intuition remains very important in making the right diagnosis per patient. So it is important for doctors to keep checking A.I. diagnoses. This requires other skills, such as a certain level of technological literacy. This means doctors need to invest time in continuing education, as technology is constantly changing. So it is doubtful whether doctors would find it feasible to take on the role of data analyst as well. Moreover, this does not appear to be realistic: the average internet user, for instance, does not know exactly what they are consenting to when they accept 'cookies'. Similarly, many doctors may not understand how an algorithm works. In a busy job, they also do not always have the time or mental space to learn to understand this, nor to evaluate the system.

As a result, 'human in the loop' control, in practice, regularly turns out to be a façade in practice.

Casus 1: who actually thought that up?

In a hospital research laboratory, dozens of doctors work passionately to diagnose serious, intergenerationally transmitted diseases. In doing so, they are helped by algorithms. A.I. recognises patterns in large data pools, and on that basis can predict who is at high risk of developing serious diseases. Those individuals can then be assisted in a disease prevention programme. Enquiries reveal that doctors do not know why A.I. detects certain diseases and not others - which, in theory, could also be easily recognised by A.I. Nor does anyone know exactly how A.I. works, nor who is actually responsible for evaluating the algorithmic outcomes. A process analysis offers some clarity: it shows that the code was built by a Programming PhD student whose PhD project was paid for by a well-known pharmaceutical company. Not coincidentally, that company makes drugs for the disease that the A.I. detects.

This case shows two things: first, that getting this code built was a political (or perhaps even commercial) choice, which is not always transparent to doctors in everyday use. Second, the case underlines that doctors, in their busy, day-to-day work, are far from always aware of the political background or the exact workings of an algorithm. Indeed, in the case of complex, self-learning A.I., nobody understands exactly how the algorithm works, not even the programmers. So an expectation that doctors should do so is unrealistic and unfair. At most, it is feasible that they learn to be critical and inquisitive towards algorithms, that standard checks are performed, and that the background of choices in programming should be transparent.

Scenario 1

The adoption of tasks and processes by A.I. requires significant adjustments from both medics and patients and training. Training should focus not only on technology, but also on the balance between technology and human interaction. Stakeholders agree that maintaining the human element in healthcare is of great importance, especially as A.I. increasingly takes over complex tasks from humans. At the same time, the line between A.I. and human input is increasingly blurred.

A.I. is becoming increasingly human, with chatbots, for example, adopting a tone that is increasingly difficult to distinguish from a human. This leads to a grey area where trust in technology is sometimes present, and sometimes not yet. Every doctor uses an Artificial Assistant, which can deliver information at lightning speed. Faster, in fact, than the best and most experienced doctor. Whereas doctors were initially still suspicious, often double-checking whether the AA is right, they are increasingly abandoning it. No time, but more importantly, no reason - the computer seems to be faultless. Patients too have become increasingly accustomed to A.I. systems and increasingly rely on them. This applies not only to everyday situations, but certainly to healthcare. Technological innovations are literally and figuratively increasingly intertwined with the patient. Patients can walk or see again because of innovations; innovations that would not be possible without technology. And 'doctor Google' is not only faster than the doctor, but also better. And sometimes nicer, because ‘doctor Google’ is never overworked.

Yet patients generally prefer a human doctor they can approach with questions, despite all the support provided by technology. Personal interaction is and will remain irreplaceable, partly because the implementation of A.I. raises new questions and concerns. The "machine data" is not always understandable and doctors need to convert the data into understandable human language to avoid patients misinterpreting it. However, A.I. also offers opportunities for a more holistic approach to better listen to the patient's story and see the bigger picture. Intuition and "gut-feeling" are therefore two unique selling points of doctors.

Remote care brings together this personal interaction and technological innovation. Doctors are now trained to deliver care via screens, which requires adaptability, both for medical staff in training and doctors who have been working in healthcare for years. Further training is therefore an important component, but the catch is that it puts an additional burden on already busy doctors.

Casus 2: New training, new skills?

A young trainee doctor learns from an experienced teacher that she can "smell" whether a patient is (seriously) ill or not. According to the lecturer, it is a skill that needs a lot of practice: it is an example of a doctor’s intuition or a gut feeling. This can be developed by seeing many patients and tuning in well to your human instinct during those encounters. The young trainee doctor looks up to the renowned lecturer and resolves to practise this skill during fellowships and then in his daily work as a junior doctor.

But in practice, this proves difficult. Most of his training is about mastering computer systems, analysing and interpreting data sets and working with consultative algorithms. During fellowships, doctors are so busy reading data off a computer screen that they usually forget to also make conscious use of their other senses, including smell. Moreover, some of his patient contacts take place digitally. The lecturer's advice to "smell" did not take this into account!

After several years of work experience, the doctor no longer thinks about the advice to sniff - it is a skill he has not developed properly and dares not rely on. Instead, he fully trusts the algorithm, with which he works smoothly and which, as far as he knows, is flawless.

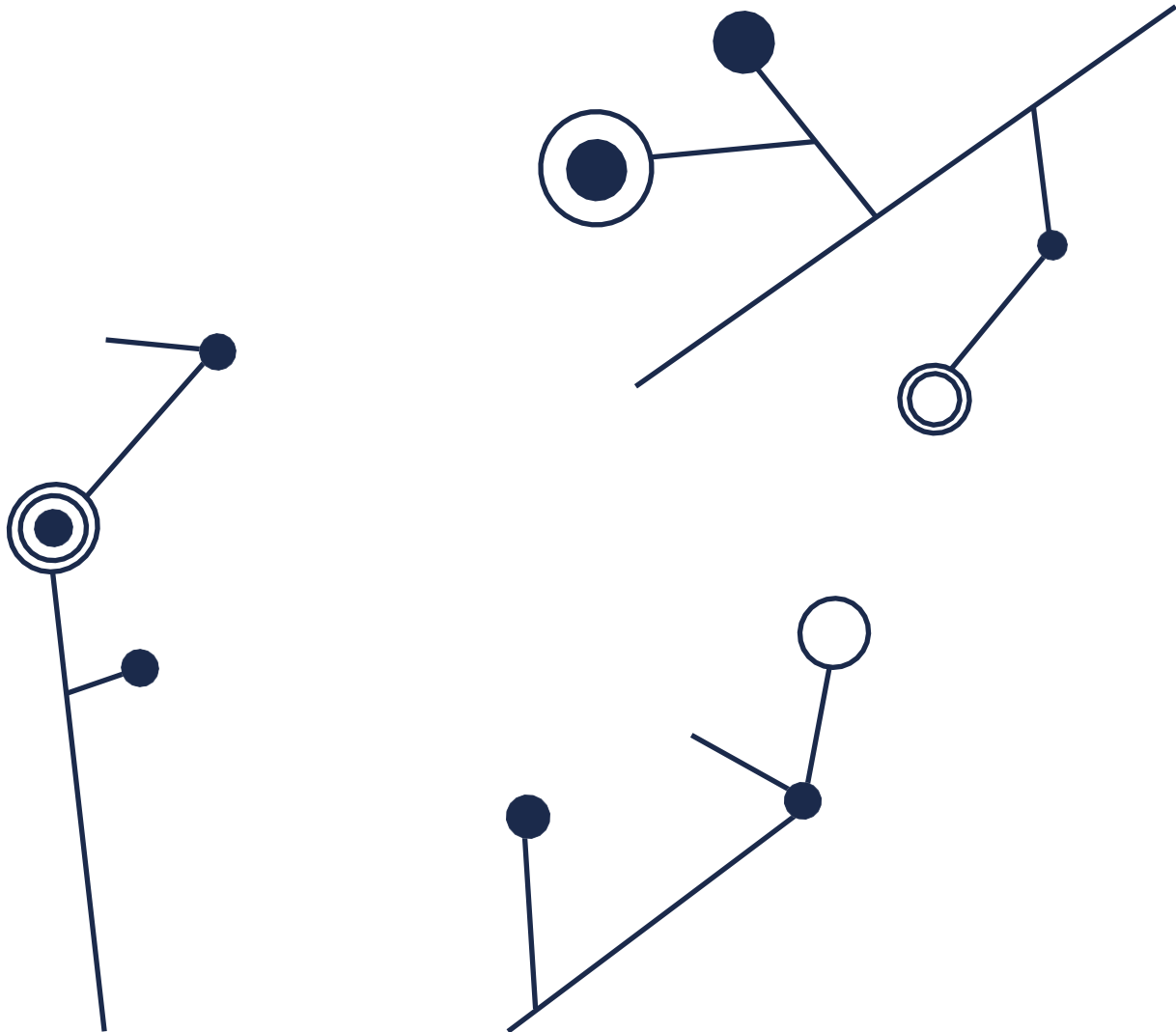
Scenario 1

Technologists and industry are increasingly collaborating and sharing broad knowledge about A.I. This presents opportunities, but also challenges. There is a danger of commercialisation, especially with regard to ownership of privacy-sensitive data being collected. Data ownership is still a thorny issue for which regulation cannot yet keep up with the high flow of innovation. Computers make fewer mistakes than tired and overworked doctors, but an A.I. error can have large-scale effects for which responsibility is not yet clear.

It is vital that doctors continue to look at whether new technology applications add value. Ideally, working groups of doctors would make decisions on the highest-priority applications that make a positive contribution.

Rapid adoption of technological innovations leads to higher risks and more chances of carelessness. Taking too low a risk, however, increases the workload to correctly train algorithms on the back end. The medical profession is increasingly changing towards collecting and monitoring data in addition to performing medical procedures. This is partly enabled by the shift from reactive to preventive healthcare. As a result, the credo "prevention is better than cure" is becoming increasingly important, and patients are taking more responsibility for monitoring their own health and acting on that data.

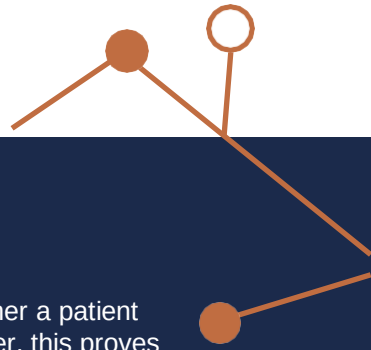
The change in the profession of physicians creates a new division of time and attention. The human touch is an irreplaceable advantage of the physical physician and remains an important and central competence. A holistic approach and customisation become more easily attainable because of the profound level of knowledge that A.I. brings.



Scenario 2 Supporting human work

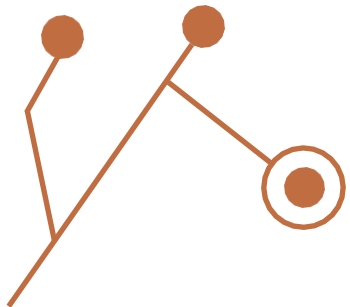
In 2033, A.I. will be used as a supporting tool to optimise the healthcare of the future. Human knowledge remains key and central.

Ensuring the "humanness" of a human doctor is considered essential, both by training courses and medical centres in which medics work. Human competences are emphasised and doctors are taught what algorithms can and cannot do, how intelligent A.I. is and where humans still outperform computers. While A.I. in the field focuses on routine tasks, doctors focus holistically on the patient. There is a focus on eye contact, open questions and using human senses during patient encounters. Since A.I. acts as a new right hand, more time is available for these things. For example, the technology helps update records and takes the initial standard contact procedures during consultations out of the hands of medics. Technical assistants perform routine checks on algorithmic outcomes, and regular reviews of the technology take place within medical teams.



Casus 3: an algorithm as a colleague

Nurses in a hospital are not allowed by protocol to independently determine whether a patient needs treatment or not. They must always seek advice from a doctor first. However, this proves difficult because the doctor is overstaffed and often arrives at the patient too late, making the nurse's advice obsolete. Moreover, nurses have more contact with their patients, allowing them to give a more complete picture. An algorithm offers a solution to this problem. Instead of the doctor giving a second opinion, the algorithm now does so. Nurses walk around the ward with iPads. For each patient, they put their own estimate next to the algorithm's data. At a glance, they see, through eye-catching colours, whether the algorithm agrees with their advice. If the algorithm agrees, they can implement it immediately. If not, they can still ask a doctor to visit. This way of working is faster and more effective, and there are fewer errors in assessment, as measured by the hospital. Nurses report feeling empowered in their work by this way of working.



Scenario 2

For now, A.I. is being implemented cautiously. A low-risk approach has been adopted to start implementing A.I. safely. This means that overly opaque or unpredictable A.I. should not be implemented as a matter of principle. Even for the simpler or imitable A.I., the doctor's intuition remains decisive in many cases. This is what junior doctors are trained in; the prevailing mantra that every student knows by heart is: 'First doctor, then A.I".

That is, A.I. provides data and results, but a doctor's intuition cannot be generated by a computer system. Therefore, doctors in training are taught that, as much as possible, they should first diagnose themselves, after which a computer system is asked to make a parallel diagnosis. The comparison is weighed - if in doubt, or if doctor and A.I appear to disagree, the doctor runs through the algorithm's decision-making process, possibly with the help of a technical assistant. Another standard rule is that if doctor and A.I disagree, a human colleague provides additional input. This is often done online, using an online, living database of doctors, in order to work together time-efficiently and neutrally.

The above shows that A.I. does a lot of shadow work, as all data has to be checked by doctors and technicians to minimise risk. This takes a lot of time. In addition, doctors spend a lot of time educating patients on how to use technology, such as home meters, alarm systems and electronic records. Also, for now, the focus is mainly on building the necessary trust in technology step by step, both from patients and doctors. Each new technological step is thus closely scrutinised by technicians, explained to doctors and patients in understandable jargon wherever possible, and responsibility is tightly framed legally.

There is regular evaluation in hospitals of what A.I. and computer systems are doing, why, and what has changed - in the case of self-learning systems - from previous use. This prevents A.I. from making wrong conclusions and diagnoses for a long time, which then generate skewed data.

However, this caution has procedural consequences. It slows down innovation and - to the frustration of many - leaves promising opportunities on hold. It also requires a lot of time to test and adjust. However - proponents say - this does produce algorithms that will perform better with each test, reducing risks. Because one thing is clear: a computer makes fewer mistakes, but íf it goes wrong, large-scale problems can arise.

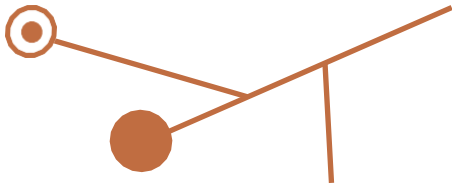
A major sticking point does remain: who is responsible if A.I. makes a mistake?

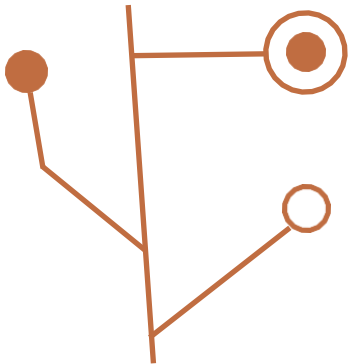
Scenario 2

Casus 4: the one zero too many

The senior doctor remembers it well: when she had just joined the department, colleagues' illegible handwriting drove her crazy. From their chicken scratch, she had to decipher what their diagnosis was, what they had noticed about the patient and what they were prescribing. More than once she was mistaken in what she thought she was reading. Then the advice seemed incomprehensible and she had to inquire with the relevant doctor. This was very inconvenient and time-consuming, especially if such a doctor had already gone home or was not to be disturbed during an operation. Fortunately, there is now a computer system in which all doctors work. In it, they digitally enter text and numbers in the appropriate tables. This is always legible, so confusion is a thing of the past. Something did go wrong recently: a doctor turned out to have entered a zero too many when writing out a medication recommendation. Nobody, including the elderly doctor, had recognised the mistake. The computer screen was so full of numbers that the extra zero went unnoticed. Even at the pharmacy, where the advice was automatically sent, no alarm was raised: after all, the pharmacist did not know the patient and her usual dose. Fortunately, the error was discovered by the patient herself, who wondered why this time she was prescribed so much more medication than usual. When she put this question to the doctor, the doctor was shocked: this dose could have been very detrimental to the patient. That same afternoon, the doctor wrote a cautionary e-mail to her colleagues, with this example in it: who is actually responsible for checking a computer screen: the doctor who fills it in, the doctor who looks at it, the pharmacist? And who gets the blame if things go wrong? And what about when A.I. itself prescribes the diagnosis, perhaps with a zero or two too many?

Both medics and patients are being prepared step by step for greater adoption of A.I. For some, this will be easier than for others. Some groups, such as the elderly, struggle to keep up with technological changes. Building trust around A.I. is also socially broadly important, as there are plenty of people who are sceptical of the latest technological innovations. This is therefore where a major challenge lies. Nevertheless, the trust around A.I. is slowly but surely growing.





Scenario 2

Casus 5: my computer, my buddy

Nurses were apprehensive when they heard that a computer was to take over some of their tasks. From now on, patients would no longer have to come to the hospital for follow-up consultations after surgery: they would have to fill in a digital form daily with questions about how they felt and how much pain they were in. An algorithm would sound an alarm if the data indicated that a patient was not doing well, so medics could intervene in those cases. A nurse could then video call the patient to find out what was going on and take additional steps if necessary. In cases where the data indicated that patients had no worrisome symptoms and the algorithm did not raise an alarm, medics needed to carry out fewer unnecessary checks.

Nurses found that the system worked better than they had thought beforehand. They had feared a dehumanisation of care, where patients would feel coldly treated without human contact. But that turned out not to be true. Anthropological research showed that patients described the computers as 'buddies' and found it pleasant and convenient to interact with the device. Indeed, patients often found it easier to be honest with a computer than with a nurse. They did not want to bother the nurses because they knew how busy they were, or they did not want to be seen as complainers. Nurses also discovered that video calling allowed them to obtain a lot of meaningful information, sometimes even more than during a patient's visit to the hospital. Precisely because they could now see the patient on the screen in his or her home environment, intimacy and insight were created. For example, they could see a large wash-up in the background, indicating that the patient was struggling to keep up with housework. Or there was a family member on screen who could informally join the conversation.

One drawback of the new digital way of working, however, was that it created shadow work. Nurses spent a relatively large amount of time analysing the data and calling back or checking alarms they did not trust. So whether the system is actually more efficient remains to be seen. Nevertheless, both patients and nurses were positive about using A.I. in healthcare.

Certain standard processes are easier to perform with A.I. support, which can contribute to efficiency. The question is whether this efficiency is maintained with wider implementation of A.I. An important aspect is that the user of the technology retains control, so that possible errors are spotted in time. These checks naturally take time, so the time-saving aspect of time may again be lost. Moreover, A.I. is built by different parties, so the responsibility for control may also lie in a grey area.

Through low-risk adoption, medical staff will be familiarised with the workings of A.I. Both experienced doctors and medical students will be increasingly immersed in A.I. A balance needs to be struck between the workload and expectations associated with A.I. The question is whether it is reasonable to expect doctors to also take on some of the tasks of data analyst.

On the other hand, A.I. also offers opportunities to focus more on job satisfaction. More time is created for subject-related issues and less time is needed for process-related actions. There is more room for a holistic approach and more complex care issues. More focus can be put on the bigger picture surrounding an individual. As process-related tasks are taken out of hands, new opportunities arise to listen better to the patient.

There are some great opportunities, but certain "blind spots" can also arise if the processes adopted by A.I. are lost sight of. Cautious implementation can lead to less awareness of potentially harmful risks. When A.I. approaches standard procedures incorrectly for years, new problems can arise with major consequences. The challenge is therefore to strike the right balance between trust in A.I. and monitoring the risks.

Contact

Dr. R.S. Roanne van Voorst


www.anthropologyofthefuture.com


r.s.vanvoorst@uva.nl

UNIVERSITEIT VAN AMSTERDAM

This report was produced in collaboration with:

Mark Hoogendoorn, Full Professor of Artificial Intelligence, Free University of Amsterdam

Carl Moons, Professor of Clinical Epidemiology and Director of Research, Julius Center for Health Sciences and Primary Care, UMC Utrecht

Kyon Soons, Ethics Advisor, a.s.r. insurance

Dieuwertje Luitse, PhD-Candidate Ethical Health & AI, University of Amsterdam

Fenna Mutsters, Head of Business Development Siilo Advisor Healthcare Policy & Management

Carine van Oosteren, Senior innovation policy advisor at TNO Vector

Leendert van Maanen, Associate Professor, Utrecht University

Peter Joosten, Future Researcher

Roy van den Heuvel, PhD-Candidate Systemic Change, Department of Industrial Design, Eindhoven University of Technology

Maarten Stol, Chief Scientific Officer, BrainCreators

Marlies Wakkee, Dermatologist, Erasmus MC Cancer Institute

Bart-Jan Verhoeff, Internist-nephrologist, St Jansdal hospital

Bas Roukema, ENT specialist, Tergooi MC

Pieter Jeekel, Chairman Dutch AI Coalition Well-being & Healthcare

Jacobien Oosterhoff MD-PhD AI for Healthcare Systems, TU Delft

Gerbert Kaandorp, Co-Founder & Board Member BrainCreators

Edith van de Weg, Senior Transformation Advisor & Partner, Science Park Inc

Nicky Hekster, Executive Professor, TIAS School for Business & Society

Michaël Bloos, Creative producer and creator, Company New Heroes

Roderick van Leerdam, Director Medical Strategy & Innovation, Philips

Frank Wartena, Program Manager Data & AI, Philips Chief Architect Office

Ilse van der Linden, PhD-Candidate XAI, Civic AI Lab, University of Amsterdam